

Anàlisi de dades mitjançant Business Intelligence: Un estudi de casos pràctics

Cristina Soler Bigorra

1 INTRODUCCIÓ - CONTEXT DEL TREBALL

El Business Intelligence (BI) consisteix en recopilar, analitzar i visualitzar dades per extreure'n conclusions i facilitar la presa de decisions. A través de diverses tecnologies, permet integrar i processar grans volums d'informació, identificar patrons, detectar anomalies i fins i tot predir tendències futures.

Un exemple clar d'ús del BI és el sector de l'entreteniment. Plataformes com Netflix, Amazon i Spotify l'utilitzen per analitzar les preferències dels usuaris i oferir recomanacions personalitzades. Per exemple, un usuari que compra a Amazon li poden recomenar productes similars als que ha estat mirant anteriorment, així augmentant les possibilitats de que es realitzi una compra. A més, no només s'aplica en aquest àmbit, sinó que també en sectors com els supermercats o les botigues de roba, on ajuda a optimitzar estocs, personalitzar ofertes i millorar l'experiència del client.

Aquest Treball de Final de Grau (TFG) explorarà el concepte i les tecnologies del BI a través d'un cas pràctic. Per començar, s'analitzaran els tiquets de compra d'un únic client en un supermercat. A partir d'aquestes dades, s'utilitzarà Power BI, una eina de Microsoft per a la visualització de dades, amb l'objectiu d'identificar patrons de consum i extraure conclusions.

Un cop completada aquesta fase inicial, el treball s'ampliarà amb un conjunt de dades més extens per aprofundir en les capacitats del BI. Això permetrà demostrar la importància del volum de dades en els processos analítics i confirmar com l'aplicació del BI pot ser un factor clau per a la personalització de serveis i la fidelització de clients en diferents sectors.

2 OBJECTIUS

L'objectiu d'aquest treball és principalment demostrar d'una forma pràctica la utilitat de la Business Intelligence. Bàsicament, veure de diferents maneres com aquesta pot ser útil tant com per un recull relativament petit de dades

en ser un cas més casolà, com per obtenir decisions a partir d'una quantitat massiva de dades.

El treball tindrà com a objectiu primer analitzar un cas amb menys quantitat de dades, per així familiaritzar-se amb el Business Intelligence, i també tindrà una segona part on s'analitzarà un conjunt més gran de les dades, veient fins a quin nivell es pot analitzar amb aquesta eina. En aquest document m'he referit a aquestes parts com a primer cas i segon cas.

Es pot veure un exemple dels tiquets analitzats a l'apèndix A1. Exemple dades primer cas.

Uns objectius més concrets serien els següents:

Identificar els patrons de les dades: buscar patrons dins de les dades, en el primer cas dels tiquets podrien ser els horaris de compra o els productes més buscats i tendències del client que s'està analitzant.

Desenvolupar una visualització de dades: utilitzar Power BI per crear informes i dashboards que permetin visualitzar les dades de manera clara i eficaç.

Extreure conclusions: a partir de les anàlisis elaborades, extreure conclusions com podrien ser, per exemple, recomanacions al client en el primer cas dels tiquets de compra.

Realitzar prediccions: usar models d'Intel·ligència Artificial per poder fer prediccions partint de les dades.

2 METODOLOGIA

La metodologia que es farà servir estarà basada en Scrum. Tot i que aquesta està dissenyada per a equips de treball, s'adaptarà i s'agafaran alguns dels seus conceptes per poder realitzar una versió individual. Així, es seguiran alguns dels seus principis per estructurar les tasques i posar fites.

El temps de treball es dividirà en Sprints, tenint en compte que el treball s'ha començat la setmana del 24 de febrer i acaba el 30 de juny. Per tant, consistirà de 18 setmanes de realització del TFG, y aquestes estaran dividides en Sprints de 3 setmanes. Per tant, el treball es farà en 6 Sprints.

Primerament, es farà un recull de les tasques a realitzar, l'equivalent en Scrum del Product Backlog. Algunes d'aquestes seran la preparació de les dades o la creació d'un dashboard en Power BI. Aquestes tasques formaran la base de la planificació del meu treball. Un cop especificades aquestes tasques, es decidirà quines d'aquestes pertanyen al Sprint actual.

Al principi del Sprint, es revisaran les tasques que s'han de complir i quina prioritat tenen. Al ser un treball individual, no hi haurà reunions d'equip, sinó que cada setmana dedicaré mitja hora per mi mateixa a revisar el que he realitzat i el que falta encara per fer i quins problemes m'han sorgit. Aquesta serà una oportunitat més per ajustar la planificació al progrés de treball que es porti. Per acabar el Sprint, també es farà una reflexió similar, analitzant si he acabat el treball que m'havia proposat i si els resultats són adients o s'hauria de fer alguna cosa més per aconseguir-ho.

A més, cal especificar que es farà servir l'eina Jira per tenir un control del procés del treball. Allà s'organitzaran i s'actualitzara el progrés de les tasques, podent així monitorar l'estat del treball de final de recerca.

3 PLANIFICACIÓ

A continuació es mostra amb detall la planificació inicial, tot i que pot estar subjecte a canvis a mesura que progressa el treball, en cas de modificacions temporals depenent del seu progrés o si el treball acaba afegint noves funcionalitats i fent-se més extens.

La planificació estarà dividida en 6 Sprints, començant el 24 de febrer i acabant el 30 de juny. Els Sprints es poden veure a continuació a la taula 1.

Sprint	Durada	Objectius
Sprint 1	24 febrer - 16 març	Planificació de l'abast del treball

		<p>Establir els requisits i tasques a realitzar</p> <p>Primer tractament de les dades del primer cas</p> <p>Primeres visualitzacions de les dades del primer cas</p>
Sprint 2	17 març - 6 abril	<p>Anàlisi del primer cas</p> <p>Refinar visualitzacions de les dades</p> <p>Extreure conclusions del primer cas</p> <p>Revisar el primer cas</p>
Sprint 3	7 abril - 27 abril	<p>Plantejar el segon cas</p> <p>Obtenció de dades del segon cas</p> <p>Tractament de dades segon cas</p>
Sprint 4	28 abril - 18 maig	<p>Analitzar dades segon cas</p> <p>Realitzar prediccions del segon cas</p> <p>Primeres visualitzacions de dades del segon cas</p>
Sprint 5	18 maig - 8 juny	<p>Conclusions segon cas</p> <p>Mostra de les dades del segon cas</p>
Sprint 6	9 juny - 30 juny	Realitzar tasques

		endarrerides
		Revisar el treball
		Ajustos finals
		Escriure entrega final del TFG i la seva presentació

Taula 1: Planificació Sprints

A més, a la imatge 1 es pot veure el diagrama de Gantt dels Sprints.



Imatge 1: Diagrama de Gantt

BIBLIOGRAFIA

- [1] "Introducción a Jira: guía completa para principiantes". Atlassian. Accedido el 3 de marzo de 2025. [En línea]. Disponible: <https://www.atlassian.com/es/software/jira/guides/getting-started/introduction#what-is-jira-software>
- [2] "Introducción | Microsoft Power BI". Accedido el 3 de marzo de 2025. [En línea]. Disponible: <https://www.microsoft.com/es-es/power-platform/products/power-bi/getting-started-with-power-bi>
- [3] R. Scrum, Agile Project Management: The Ultimate Step by Step Guide to Learn Agile Project Management to Complete Your Goals with Maximum of Results. Independently Publ., 2019.
- [4] Kevin Stratvert. How to use Microsoft Power BI - Tutorial for Beginners. (4 de agosto de 2020). Accedido el 3 de marzo de 2025. [Video en línea]. Disponible: https://www.youtube.com/watch?v=TmhQCQr_DC_A
- [5] "Implementación de Soluciones de Business Intelligence (BI) con Python". 10Code Software-Entwicklung. Accedido el 3 de marzo de 2025. [En línea]. Disponible: <https://10code.es/de/bi-python/>
- [6] "W3Schools.com". W3Schools Online Web Tutorials. Accedido el 3 de marzo de 2025. [En línea]. Disponible: <https://www.w3schools.com/python/pandas/default.asp>
- [7] "W3Schools.com". W3Schools Online Web Tutorials. Accedido el 3 de marzo de 2025. [En línea]. Disponible: <https://www.w3schools.com/python/default.asp>
- [8] "What are the phases/stages of a business intelligence project?" Le blog des meilleurs consultants indépendants I FocusTribes. Accedido el 3 de marzo de 2025. [En línea]. Disponible: <https://blog.focustribes.com/en/business-intelligence-phases>
- [9] "A Gantt Chart Guide with Definitions & Examples". ProjectManager. Accedido el 3 de marzo de 2025. [En línea]. Disponible: <https://www.projectmanager.com/guides/gantt-chart>
- [10] D. Lopez. "Cómo extraer datos de archivos PDF con Python". freeCodeCamp.org. Accedido el 3 de marzo de 2025. [En línea]. Disponible: <https://www.freecodecamp.org/espanol/news/como-extraer-datos-de-archivos-en-pdf/>
- [11] freeCodeCamp.org. Machine Learning for Everybody – Full Course. (26 de septiembre de 2022). Accedido el 3 de marzo de 2025. [Video en línea]. Disponible: https://www.youtube.com/watch?v=i_LwzRVP7bg

APÈNDIX

A1. EXEMPLE DADES PRIMER CAS



MERCADONA, S.A. A-46103834

C/ VILADOMAT 195

08205 SABADELL

TELÈFON: 937208041

11/02/2025 19:17 OP: 1656930

FACTURA SIMPLIFICADA: 3441-015-803829



Descripció	P. Unit	Import
2 LLET SENCERA	0,94	1,88
1 BEGUDA AVENA		0,95
1 BOSSA PLÀSTIC		0,15
4 PA VIENA	0,40	1,60
1 RAÏM VERMELLA SESE		2,69
1 MAGDALENA SG 6UDS		2,55
1 ENCENEDOR CUINA WAVE		2,57
1 CUIDACOL NATURAL		2,46
1 BÍF FRUITES SIL		1,23
1 FORMATGE TALLS LIGHT		2,46
1 MANDARINA		
1,404 kg	2,19 €/kg	3,07
1 PÀRQUING		0,00
ENTRADA 19:02 SORTIDA 19:17		
TOTAL (€)		21,61
TARGETA BANCÀRIA		21,61

IVA	BASE IMPOSABLE (€)	QUOTA (€)
4%	13,26	0,53
10%	4,64	0,46
21%	2,25	0,47
TOTAL	20,15	1,46



S'ADMETEN DEVOLUCIONS AMB TIQUET

DISPOSA DE 20 MINUTS
PER RETIRAR EL SEU VEHICLE